Linear Algebra and its Applications 430 (2009) 2182–2191

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

# Inner functions of numerical contractions

Hwa-Long Gau [a,*,1], Pei Yuan Wu [b,2]

[a] Dept. of Mathematics, National Central University, Chungli 32001, Taiwan
[b] Dept. of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan

ARTICLE INFO

*Article history:*
Received 1 September 2008
Accepted 24 November 2008
Available online 30 December 2008

Submitted by A. Berman

*AMS classification:*
47A12
47B38

*Keywords:*
Numerical range
Numerical radius
Numerical contraction
Compression of the shift

ABSTRACT

We prove that, for a function $f$ in $H^\infty$ of the unit disc with $\|f\|_\infty \leqslant 1$, the existence of an operator $T$ on a complex Hilbert space $H$ with its numerical radius at most one and with $\|f(T)x\| = 2$ for some unit vector $x$ in $H$ is equivalent to that $f$ be an inner function with $f(0) = 0$. This confirms a conjecture of Drury [S.W. Drury, Symbolic calculus of operators with unit numerical radius, Linear Algebra Appl. 428 (2008) 2061–2069]. Moreover, we also show that any operator $T$ satisfying the above conditions has a direct summand similar to the compression of the shift $S(\phi)$, where $\phi(z) = zf(z)$ for $|z| < 1$. This generalizes the result of Williams and Crimmins [J.P. Williams, T. Crimmins, On the numerical radius of a linear operator, Amer. Math. Monthly 74 (1967) 832–833] for $f(z) = z$ and of Crabb [M.J. Crabb, The powers of an operator of numerical radius one, Michigan Math. J. 18 (1971) 253–256] for $f(z) = z^n$ $(n \geqslant 2)$.



For a bounded linear operator $A$ on a complex Hilbert space $H$, its *numerical range* and *numerical radius* are

$$W(A) = \{\langle Ax, x\rangle : x \in H, \|x\| = 1\}$$

and

$$w(A) = \sup\{|z| : z \in W(A)\},$$

respectively, where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ denote the inner product and its associated norm in $H$. It is known that $W(A)$ is a bounded convex subset of the plane. When $H$ is finite dimensional, it is even compact.

* Corresponding author.
*E-mail addresses:* hlgau@math.ncu.edu.tw (H.-L. Gau), pywu@math.nctu.edu.tw (P.Y. Wu).

1 Research supported by the National Science Council of the Republic of China under NSC 96-2115-M-008-006.
2 Research supported by the National Science Council of the Republic of China under NSC 96-2115-M-009-013-MY3 and by the MOE-ATU.

0024-3795/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2008.11.020

Its closure $\overline{W(A)}$ contains the spectrum $\sigma(A)$ of $A$. For other properties of the numerical range and numerical radius, the reader may consult [11, Chapter 22] or [10].

An operator $A$ is a *numerical contraction* (resp., *contraction*) if $w(A) \leqslant 1$ (resp., $\|A\| \leqslant 1$). In 1967, Sz.-Nagy and Foiaş [16] proved that every numerical contraction is similar to a contraction. Some years later, Okubo and Ando [13] gave another proof basing it on a factorization of the numerical contraction by Ando [1], which has the advantage of a sharp control on the invertible operator implementing the similarity. As a consequence, an estimate on the norm of a function of a numerical contraction can easily be obtained.

**Theorem 1.** (a) *An operator $A$ is a numerical contraction if and only if $A = 2(I - B^*B)^{1/2}B$ for some contraction $B$.*

(b) *If $A$ is a numerical contraction, then $A = XCX^{-1}$ for some invertible operator $X$ with $\|X\|, \|X^{-1}\| \leqslant \sqrt{2}$ and some contraction $C$.*

(c) *If $A$ is a numerical contraction and $f : \overline{\mathbb{D}} \to \mathbb{C}$ is a function analytic on $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and continuous on $\overline{\mathbb{D}}$, then $\|f(A)\| \leqslant 2\|f\|_\infty$, where $\|f\|_\infty = \sup\{|f(z)| : z \in \overline{\mathbb{D}}\}$.*

For our later use, we briefly sketch a proof of Theorem 1(b) based on (a), which is slightly different from the one in [13, Theorem 2]. Let $A$ be factored as in (a). If

$$g(t) = \begin{cases} \sqrt{2(1-t)} & \text{if } 0 \leqslant t \leqslant 1/2, \\ 1/\sqrt{2t} & \text{if } 1/2 \leqslant t \leqslant 1, \end{cases} \tag{1}$$

then both $g$ and $1/g$ are continuous functions on $[0, 1]$ with $\|g\|_\infty = \|1/g\|_\infty = \sqrt{2}$, where $\|\cdot\|_\infty$ denotes the supremum of a function over $[0, 1]$. It is easily seen that $X \equiv g(B^*B)$ is invertible, $\|X\|, \|X^{-1}\| \leqslant \sqrt{2}$ and

$$\begin{aligned} \|X^{-1}AX\| &\leqslant 2\|g(B^*B)^{-1}(I - B^*B)^{1/2}\| \cdot \|(B^*B)^{1/2}g(B^*B)\| \\ &\leqslant 2 \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = 1. \end{aligned}$$

More recently, Drury [7] in studying the norm and numerical radius of $f(A)$ proposed a conjecture on the sharpness of the inequality in Theorem 1(c). The purpose of this paper is to confirm this conjecture with a more detailed information on the structure of $A$.

In the following, we will consider a more general functional calculus than the one in Theorem 1(c) for numerical contractions. Indeed, if $A$ is a numerical contraction on $H$, then the Berger dilation theorem [3] says that there is a unitary operator $U$ on a space $K$ containing $H$ such that $A^n = 2P_HU^n|H$ for all $n \geqslant 1$, where $P_H$ denotes the (orthogonal) projection from $K$ onto $H$. Such a unitary 2-dilation $U$ of $A$ can be taken to be minimal in the sense that $K = \bigvee\{U^nH : n = 0, \pm 1, \pm 2, \ldots\}$. In this case, $U$ is uniquely determined up to isomorphism, and, moreover, if $A$ is completely nonunitary, that is, if $A$ has no unitary direct summand, then $U$ is absolutely continuous (cf. [8, Theorem 1] and [14, Proposition 2]). (We thank G. Cassier and L. Kérchy for providing us the relevant references on this subject.) Hence if $A' = U' \oplus A$ on $L \oplus H$ is a numerical contraction, where $U'$ is absolutely continuous unitary and $A$ is completely nonunitary, then $f(A') \equiv f(U') \oplus ((2P_Hf(U)|H) - f(0)I)$ for $f$ in $H^\infty$ is well-defined, where $U$ is the minimal unitary 2-dilation of $A$. Note that Theorem 1(c) is obviously true for $A$ a numerical contraction with no singular unitary part and $f$ in $H^\infty$.

For an inner function $\phi$ ($\phi$ bounded analytic on $\mathbb{D}$ with $|\phi| = 1$ almost everywhere on $\partial\mathbb{D}$), the *compression of the shift* $S(\phi)$ is defined on $H(\phi) = H^2 \ominus \phi H^2$ by

$$S(\phi)f = P_{H(\phi)}(zf(z))|H(\phi) \quad \text{for } f \in H(\phi).$$

Such operators have been studied extensively since the 1960s starting with the work of Sarason [15]. A nice account of their properties together with those of the more general $C_0$ contractions can be found in [2]. Sz.-Nagy and Foiaş [17] is the classical treatise on further developments of this subject. In particular, if $\phi$ is a Blaschke product with $n$ zeros (counting multiplicity), then $H(\phi)$ is $n$-dimensional.

Our main result is the following:

**Theorem 2.** *Let $f$ be a function in $H^\infty$ with $\|f\|_\infty \leqslant 1$. Then there exists a numerical contraction $T$ with no unitary part such that $\|f(T)x\| = 2$ for some unit vector $x$ if and only if $f$ is inner and $f(0) = 0$. Moreover, any operator $T$ satisfying the above conditions has a direct summand similar to $S(\phi)$, where $\phi(z) = zf(z)$ for $|z| < 1$.*

A finite-dimensional version of this confirms Drury's Conjecture 6 in [7].

**Corollary 3.** *Let $f : \overline{\mathbb{D}} \to \mathbb{C}$ be analytic on $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$ with $\|f\|_\infty \leqslant 1$. Then there exists a numerical contraction $T$ with $\|f(T)x\| = 2$ for some unit vector $x$ if and only if $f$ is a finite Blaschke product and $f(0) = 0$. In this case, if $f$ has $n$ zeros (counting multiplicity), then any such $T$ is unitarily equivalent to an operator of the form $A \oplus A'$, where $A$ can be represented by the $(n+1)$-by-$(n+1)$ upper-triangular matrix $[a_{ij}]_{i,j=1}^{n+1}$ with $a_i \equiv a_{ii}$ satisfying $a_1 = a_{n+1} = 0$ and $|a_i| < 1$ for all $i$, and*

$$a_{ij} = \begin{cases} \sqrt{2}b_{ij} & \text{if} \quad 1 = i < j \leqslant n \text{ or } 2 \leqslant i < j = n+1, \\ 2b_{ij} & \text{if} \quad i = 1 \text{ and } j = n+1, \\ b_{ij} & \text{if} \quad 2 \leqslant i < j \leqslant n, \\ 0 & \text{if} \quad i > j, \end{cases}$$

*where*

$$b_{ij} = (-1)^{j-i-1}\bar{a}_{i+1}\cdots\bar{a}_{j-1}[(1-|a_i|^2)(1-|a_j|^2)]^{1/2} \quad \text{for } i < j.$$

The matrix form of $A$ here is a consequence of Theorem 8(b) below and the matrix representation of the finite-dimensional compression of the shift $S(\phi)$ (cf. [9, Corollary 1.3]).
A special case of this yields a result of Crabb [5, Theorem 2].

**Corollary 4.** *If $T$ is a numerical contraction and $\|T^n x\| = 2$ for some $n \geqslant 1$ and some unit vector $x$, then $T$ is unitarily equivalent to an operator of the form $A \oplus A'$, where $A$ is the $(n+1)$-by-$(n+1)$ matrix*

$$\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & \sqrt{2} & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & 1 & \\ & & & & 0 & \sqrt{2} \\ & & & & & 0 \end{bmatrix}$$

*depending on whether $n = 1$ or $n \geqslant 2$.*

The case $n = 1$ was obtained earlier by Williams and Crimmins [18]. It will be invoked in the proof of Theorem 8(b).
We start by proving the sufficiency part of Theorem 2.

**Theorem 5.** *Let $f$ be an inner function with $f(0) = 0$ and let $\phi(z) = zf(z)$ for $|z| < 1$. Let $X = \sqrt{2} \oplus I \oplus (1/\sqrt{2})$ on $H(\phi) = H_1 \oplus H_2 \oplus H_3$, where $H_1 = \ker S(\phi)$, $H_3 = \ker S(\phi)^*$ and $H_2 = H(\phi) \ominus (H_1 \oplus H_3)$, and let $A = XS(\phi)X^{-1}$. Then $A$ is a cyclic irreducible operator with no unitary part such that $\overline{W(A)} = \overline{\mathbb{D}}$ and $\|f(A)x\| = 2$ for some unit vector $x$.*

The next corollary is a special case (cf. [6, Theorem 3.1]).

**Corollary 6.** *If $f$ is a Blaschke product with $n$ zeros (counting multiplicity), then there is an $(n+1)$-by-$(n+1)$ matrix $A$ with $W(A) = \overline{\mathbb{D}}$ and $\|f(A)\| = 2$.*

An operator $A$ on $H$ is *cyclic* with *cyclic vector* $x$ if $H = \bigvee\{A^n x : n \geqslant 0\}$. It is easily seen that for a cyclic $A$ the dimension of $\ker A^*$ is at most one.

An operator is *irreducible* if it is not unitarily equivalent to the direct sum of two other operators. To prove the irreducibility of the operator $A$ in Theorem 5, we need the following lemma.

**Lemma 7.** *If $A$ is cyclic with a cyclic vector in* $\ker A^*$, *then $A$ is irreducible*.

**Proof.** Assume that $A = A_1 \oplus A_2$ on $H = H_1 \oplus H_2$. Let $x = x_1 \oplus x_2$, where $x_j \in H_j, j = 1, 2$, be a cyclic vector of $A$ in $\ker A^*$. Then $A^*x = (A_1 \oplus A_2)^*(x_1 \oplus x_2) = 0$ implies that $(A_1 \oplus A_2)^*(x_1 \oplus 0) = (A_1 \oplus A_2)^*(0 \oplus x_2) = 0$. On the other hand, since $H_1 \oplus H_2 = \bigvee\{A_1^n x_1 \oplus A_2^n x_2 : n \geqslant 0\}$, we infer that $x_j \neq 0$ for $j = 1, 2$. Thus $x_1 \oplus 0$ and $0 \oplus x_2$ are linearly independent, and, therefore, $\dim \ker A^* = \dim \ker(A_1 \oplus A_2)^* \geqslant 2$, a contradiction. This proves our assertion. □

**Proof of Theorem 5.** Since $\phi(0) = 0$, the function $g \equiv 1$ is in $H(\phi)$. It is a unit cyclic vector for $S(\phi)$ and generates the one-dimensional subspace $H_3$. On the other hand, from the facts that $f$ is inner and $\phi(z) = zf(z)$ on $\mathbb{D}$ we can easily check that $f = P_{H(\phi)}f = f(S(\phi))g$ and $S(\phi)f = 0$. Thus $f$ is a unit vector which generates the one-dimensional $H_1$. That $f$ and $g$ are orthogonal follows from a simple computation using $f(0) = 0$. Note also that

$$f(A)g = Xf(S(\phi))X^{-1}g = \sqrt{2}Xf(S(\phi))g = \sqrt{2}Xf = 2f,$$

which shows that $\|f(A)g\| = 2$. Since $g \in \ker S(\phi)^*$ is a cyclic vector for $S(\phi)$, $Xg = g/\sqrt{2} \in \ker A^*$ is cyclic for $A = XS(\phi)X^{-1}$. The irreducibility of $A$ then follows from Lemma 7. Moreover, since $S(\phi)^n$ converges to 0 in the strong operator topology (SOT), the same is true for $A^n$. Hence $A$ has no unitary part.

To prove that $\overline{W(A)} \subseteq \overline{\mathbb{D}}$, let $B = S(\phi)X^{-1}/\sqrt{2}$. Since $\operatorname{rank}(I - S(\phi)^*S(\phi)) = 1$ and $S(\phi)^*S(\phi)f = 0$, we have $S(\phi)^*S(\phi) = 0 \oplus I \oplus 1$ and hence $B^*B = 0 \oplus (1/2)I \oplus 1$ on $H(\phi) = H_1 \oplus H_2 \oplus H_3$. Therefore, $B$ is a contraction and

$$\begin{aligned}2(I - B^*B)^{1/2}B &= 2\left(1 \oplus \frac{1}{\sqrt{2}}I \oplus 0\right)\frac{1}{\sqrt{2}}S(\phi)X^{-1}\\ &= XS(\phi)X^{-1} = A.\end{aligned}$$

Theorem 1(a) then implies that $\overline{W(A)} \subseteq \overline{\mathbb{D}}$.

To prove the converse, let $\lambda$ be any point in $\mathbb{D}$. Then the operator $I - \bar{\lambda}S(\phi)$ is invertible and $u \equiv (I - \bar{\lambda}S(\phi))^{-1}g - g = \sum_{n=1}^{\infty}(\bar{\lambda}S(\phi))^n g$ in norm. Let $v = u - \langle u, f\rangle f$. Note that

$$\begin{aligned}\langle v, g\rangle &= \sum_{n=1}^{\infty}\bar{\lambda}^n\langle S(\phi)^n g, g\rangle - \langle u, f\rangle\langle f, g\rangle\\ &= 0 - \langle u, f\rangle \cdot 0 = 0\end{aligned}$$

and

$$\begin{aligned}\langle v, f\rangle &= \langle u, f\rangle - \langle u, f\rangle\langle f, f\rangle\\ &= \langle u, f\rangle - \langle u, f\rangle = 0.\end{aligned}$$

Hence $v$ is in $H_2$. Finally, letting $y = \langle u, f\rangle f \oplus \sqrt{2}v \oplus g$ in $H(\phi) = H_1 \oplus H_2 \oplus H_3$, we show that $\bar{\lambda}By = (I - B^*B)^{1/2}y$. Indeed, on the one hand, we have

$$\begin{aligned}\bar{\lambda}By &= \bar{\lambda}S(\phi)\left(\frac{1}{2} \oplus \frac{1}{\sqrt{2}}I \oplus 1\right)(\langle u, f\rangle f \oplus \sqrt{2}v \oplus g)\\ &= \bar{\lambda}\left(\frac{1}{2}\langle u, f\rangle S(\phi)f + S(\phi)v + S(\phi)g\right)\\ &= \bar{\lambda}[S(\phi)(I - \bar{\lambda}S(\phi))^{-1}g - S(\phi)g] + \bar{\lambda}S(\phi)g\\ &= \bar{\lambda}S(\phi)(I - \bar{\lambda}S(\phi))^{-1}g.\end{aligned}$$

On the other hand,

$$\begin{aligned}(I-B^*B)^{1/2}y &= \left(1\oplus\frac{1}{\sqrt{2}}I\oplus 0\right)(\langle u,f\rangle f\oplus\sqrt{2}v\oplus g)\\ &= \langle u,f\rangle f+v\\ &= u\\ &= (I-\bar{\lambda}S(\phi))^{-1}g-g\\ &= \bar{\lambda}S(\phi)(I-\bar{\lambda}S(\phi))^{-1}g.\end{aligned}$$

Thus $\bar{\lambda}By=(I-B^*B)^{1/2}y$ holds. Hence

$$|\lambda|^2\|By\|^2=\|(I-B^*B)^{1/2}y\|^2=\|y\|^2-\|By\|^2,$$

which implies that $\|By\|^2=\|y\|^2/(1+|\lambda|^2)$. Therefore,

$$\begin{aligned}\langle Ay,y\rangle &= \langle 2(I-B^*B)^{1/2}By,y\rangle\\ &= 2\langle By,(I-B^*B)^{1/2}y\rangle = 2\langle By,\bar{\lambda}By\rangle\\ &= 2\lambda\|By\|^2=\frac{2\lambda}{1+|\lambda|^2}\|y\|^2.\end{aligned}$$

This shows that $2\lambda/(1+|\lambda|^2)$ is in $W(A)$ for any $\lambda$ in $\mathbb{D}$. Hence $\mathbb{D}\subseteq W(A)$ and thus $\overline{W(A)}=\overline{\mathbb{D}}$ as asserted. This completes the proof. □

We now proceed to prove the necessity part of Theorem 2.

**Theorem 8.** *Let $f$ be a function in $H^\infty$ with $\|f\|_\infty\leqslant 1$. If $T$ is a numerical contraction with no singular unitary part such that $\|f(T)x\|=2$ for some unit vector $x$, then*

(a) *$f$ is inner with $f(0)=0$, and*
(b) *$T$ is unitarily equivalent to an operator of the form $XS(\phi)X^{-1}\oplus A'$, where $\phi(z)=zf(z)$ for $|z|<1$ and $X=\sqrt{2}\oplus I\oplus(1/\sqrt{2})$ on $H(\phi)=H_1\oplus H_2\oplus H_3$ ($H_1=\ker S(\phi)$ and $H_3=\ker S(\phi)^*$).*

For the proof of its part (b), we need the following lemma.

**Lemma 9.** *Let $A$ be a $C_0$ contraction on $H$ with minimal function $\phi$. Then there is an operator $\widetilde{A}$ on $\widetilde{H}\supseteq H$ of class $C_0$ such that* (a) *$\widetilde{A}H\subseteq H$,* (b) *$A=\widetilde{A}|H$, and* (c) *$\widetilde{A}$ is unitarily equivalent to $\sum_{n=1}^{d}\oplus S(\phi)$, where $d=\operatorname{rank}(I-A^*A)^{1/2}\leqslant\infty$.*

This appeared in [12, Lemma 4] (with $T$ there replaced by $A^*$) and is dependent on the Sz.-Nagy–Foiaş contraction theory.

**Proof of Theorem 8.** (a) That $f(0)=0$ follows from Drury [7, Theorem 4]. Indeed, since the latter is also valid for functions $f$ in $H^\infty$ with $\|f\|_\infty\leqslant 1$, we have $\|f(T)\|\leqslant\nu(|f(0)|)$, where

$$\nu(t)=(2-3t^2+2t^4+2(1-t^2)(1-t^2+t^4)^{1/2})^{1/2}\quad\text{for } 0\leqslant t\leqslant 1.$$

Our assumption yields that

$$2=\|f(T)x\|\leqslant\|f(T)\|\leqslant\nu(|f(0)|)\leqslant 2$$

or $\nu(|f(0)|)=2$. This is equivalent to $f(0)=0$.

Let $M=\bigvee\{T^nx:n\geqslant 0\}$ and $A=T|M$. Then $w(A)\leqslant 1$ and $\|f(A)x\|=\|f(T)x\|=2$. By Theorem 1(a), $A=2(I-B^*B)^{1/2}B$ for some contraction $B$. Let $g$ be as in (1) and $X=g(B^*B)$. Then, as indicated before, $X$ is positive definite and invertible with $\|X\|,\|X^{-1}\|\leqslant\sqrt{2}$ and $C\equiv X^{-1}AX$ is a contraction. It is easily seen that $C$, being similar to the operator $A$ with no singular unitary part, is itself without singular unitary part. Thus $f(C)$ is well-defined. The chain of inequalities

$$\begin{aligned}2 &= \|f(A)x\| = \|Xf(C)X^{-1}x\| \\ &\leqslant \|X\|\|f(C)X^{-1}x\| \leqslant \|X\|\|f(C)\|\|X^{-1}x\| \\ &\leqslant \|X\|\|f(C)\|\|X^{-1}\| \leqslant \sqrt{2}\|f\|_\infty\sqrt{2} \leqslant 2,\end{aligned}$$

where $\|f(C)\| \leqslant \|f\|_\infty$ is by the von Neumann inequality, yields equalities throughout. In particular, we have

$$\|X\| = \|X^{-1}\| = \|f(C)X^{-1}x\| = \|X^{-1}x\| = \sqrt{2}$$

and $\|f(C)\| = \|f\|_\infty = 1$. Note that for a positive semidefinite operator $Y$ and vector $u$, the equalities $\|Yu\| = \|Y\|\|u\|$ and $Yu = \|Y\|u$ are equivalent. Thus from $\|X^{-1}x\| = \sqrt{2} = \|X^{-1}\|\|x\|$, we infer that $X^{-1}x = \sqrt{2}x$ or $Xx = (1/\sqrt{2})x$. Similarly, for $y \equiv f(C)x$, we have

$$\|y\| = \|f(C)x\| = \frac{1}{\sqrt{2}}\|f(C)X^{-1}x\| = 1$$

and

$$\|Xy\| = \|Xf(C)x\| = \frac{1}{\sqrt{2}}\|Xf(C)X^{-1}x\| = \frac{1}{\sqrt{2}}\|f(A)x\| = \sqrt{2} = \|X\|\|y\|.$$

As above, this yields $Xy = \sqrt{2}y$. Thus $x$ and $y$ are eigenvectors associated with the eigenvalues $1/\sqrt{2}$ and $\sqrt{2}$ of the positive definite $X$, respectively. Hence they are orthogonal to each other. Since $X = g(B^*B)$ with $g$ defined in (1), we infer that 1 and 0 are eigenvalues of $B^*B$ with corresponding eigenvectors $x$ and $y$, respectively. We also have

$$f(A)x = Xf(C)X^{-1}x = \sqrt{2}Xf(C)x = \sqrt{2}Xy = 2y. \tag{2}$$

From $B^*By = 0$, we obtain $By = 0$. Thus

$$Af(A)x = 2Ay = 2(I - B^*B)^{1/2}By = 0$$

and, consequently,

$$Af(A)A^n x = A^n(Af(A)x) = 0$$

for all $n \geqslant 0$. Since $M$ is generated by $A^n x$, $n \geqslant 0$, this yields $Af(A) = 0$. Hence $Cf(C) = X^{-1}Af(A)X = 0$, which shows that $C$ is a $C_0$ contraction. Let $\psi$ be its minimal (inner) function, and let $\phi(z) = zf(z)$. Then $\psi$ divides $\phi$. We necessarily have $\psi(0) = 0$ for otherwise $\psi$ would divide $f$, which would imply $f(C) = 0$, contradicting $\|f(C)\| = 1$. Hence $\psi(z) = z\eta(z)$ for some inner function $\eta$ and $f(z) = \xi(z)\eta(z)$ for some $\xi$ in $H^\infty$ with $\|\xi\|_\infty = 1$. Let $\xi(z) = \xi(0) + z\zeta(z)$ for $\zeta$ in $H^\infty$. We have $f(z) = \xi(0)\eta(z) + \zeta(z)\psi(z)$ and thus $f(C) = \xi(0)\eta(C)$. From

$$1 = \|f(C)\| = |\xi(0)|\|\eta(c)\| \leqslant \|\eta(C)\| \leqslant 1,$$

we obtain $|\xi(0)| = 1$. Therefore, $\xi(z) = \xi(0)$ is constant and $f = \xi(0)\eta$ is inner.

(b) We first show that $C$ is unitarily equivalent to $S(\phi)$, where $\phi(z) = zf(z)$. Note that, from the proof of (a), $\phi$ is the minimal function of $C$. By Lemma 9, $C$ can be extended to (an operator unitarily equivalent to) $\sum_{n=1}^\infty \oplus S(\phi)$. Hence $f(C)$ extends to $\sum_{n=1}^\infty \oplus f(S(\phi))$. Let $x = \sum_{n=1}^\infty \oplus g_n$ with $g_n$ in $H(\phi)$ for all $n$. We infer from

$$1 = \|y\|^2 = \|f(C)x\|^2 = \sum_{n=1}^\infty \|f(S(\phi))g_n\|^2 \leqslant \sum_{n=1}^\infty \|g_n\|^2 = \|x\|^2 = 1$$

that $\|f(S(\phi))g_n\| = \|g_n\|$ for all $n$. Since $f(S(\phi))$ is a contraction, we have $f(S(\phi))^*f(S(\phi))g_n = g_n$. Thus $g_n$ is in $\operatorname{ran} f(S(\phi))^*$, a one-dimensional space generated by the function $g \equiv 1$. Hence, for each $n \geqslant 1$, $g_n = a_n g$ for some scalar $a_n$. Define the operator $V : M \to H(\phi)$ by

$$V(p(C)x) = p(S(\phi))g$$

for any polynomial $p$. Since $p(C)x = \sum_{n=1}^{\infty} \oplus p(S(\phi))g_n$, we have

$$\|p(C)x\| = \left(\sum_{n=1}^{\infty} \|p(S(\phi))g_n\|^2\right)^{1/2} = \|p(S(\phi))g\| \left(\sum_{n=1}^{\infty} |a_n|^2\right)^{1/2}$$
$$= \|p(S(\phi))g\|\|x\| = \|p(S(\phi))g\|.$$

Note that $M$ being generated by $A^n x$, $n \geqslant 0$, is also generated by $C^n X^{-1} x = \sqrt{2} C^n x$, $n \geqslant 0$. Thus the set of vectors $p(C)x$, $p$ polynomial, is dense in $M$. From above, we obtain that $V$ is an isometry with $VC = S(\phi)V$. Since $\phi$ is the minimal function of $C$, the unitary equivalence of $C$ and $S(\phi)$ follows.

Let $H_1$ and $H_3$ be the one-dimensional subspaces of $M$ which are generated by $y$ and $x$, respectively, and let $H_2 = M \ominus (H_1 \oplus H_3)$. On $M = H_1 \oplus H_2 \oplus H_3$, the operators $X$ and $B^*B$ can be decomposed as $X = \sqrt{2} \oplus X_1 \oplus (1/\sqrt{2})$ and $B^*B = 0 \oplus D \oplus 1$. Let $B = [B_{ij}]_{i,j=1}^3$ on $M = H_1 \oplus H_2 \oplus H_3$. From $B^*B = 0 \oplus D \oplus 1$, we obtain $B_{11}^*B_{11} + B_{21}^*B_{21} + B_{31}^*B_{31} = 0$, which implies that $B_{11}$, $B_{21}$ and $B_{31}$ are all zero operators. Hence

$$\begin{aligned} A &= 2(I - B^*B)^{1/2}B \\ &= 2\begin{bmatrix} 1 & & \\ & (I-D)^{1/2} & \\ & & 0 \end{bmatrix}\begin{bmatrix} 0 & B_{12} & B_{13} \\ 0 & B_{22} & B_{23} \\ 0 & B_{32} & B_{33} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2B_{12} & 2B_{13} \\ 0 & 2(I-D)^{1/2}B_{22} & 2(I-D)^{1/2}B_{23} \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{3}$$

We now show that $X_1 = I$. This is done by proving $DB_{22} = B_{22}/2$ and $DB_{23} = B_{23}/2$. Note that

$$\begin{aligned} C &= X^{-1}AX \\ &= \begin{bmatrix} 1/\sqrt{2} & & \\ & X_1^{-1} & \\ & & \sqrt{2} \end{bmatrix}\begin{bmatrix} 0 & 2B_{12} & 2B_{13} \\ 0 & 2(I-D)^{1/2}B_{22} & 2(I-D)^{1/2}B_{23} \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \sqrt{2} & & \\ & X_1 & \\ & & 1/\sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \sqrt{2}B_{12}X_1 & B_{13} \\ 0 & 2X_1^{-1}(I-D)^{1/2}B_{22}X_1 & \sqrt{2}X_1^{-1}(I-D)^{1/2}B_{23} \\ 0 & 0 & 0 \end{bmatrix} \equiv \begin{bmatrix} 0 & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Since

$$I - C^*C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & I - C_{12}^*C_{12} - C_{22}^*C_{22} & * \\ 0 & * & 1 - |C_{13}|^2 - C_{23}^*C_{23} \end{bmatrix}$$

has rank one, we have

$$C_{12}^*C_{12} + C_{22}^*C_{22} = I \tag{4}$$

and

$$|C_{13}|^2 + C_{23}^*C_{23} = 1. \tag{5}$$

From (4), we obtain

$$\begin{aligned} I &= 2X_1^*B_{12}^*B_{12}X_1 + 4X_1^*B_{22}^*(I-D)^{1/2}X_1^{*-1}X_1^{-1}(I-D)^{1/2}B_{22}X_1 \\ &= 2X_1(B_{12}^*B_{12} + 2B_{22}^*X_1^{-2}(I-D)B_{22})X_1. \end{aligned} \tag{6}$$

Note that

$$B^*B = \begin{bmatrix} 0 & 0 & 0 \\ B_{12}^* & B_{22}^* & B_{32}^* \\ B_{13}^* & B_{23}^* & B_{33}^* \end{bmatrix} \begin{bmatrix} 0 & B_{12} & B_{13} \\ 0 & B_{22} & B_{23} \\ 0 & B_{32} & B_{33} \end{bmatrix} = \begin{bmatrix} 0 & & \\ & D & \\ & & 1 \end{bmatrix}$$

yields $B_{12}^*B_{12} + B_{22}^*B_{22} + B_{32}^*B_{32} = D$. We derive from (6) that $(1/2)X_1^{-2} = D - B_{22}^*B_{22} - B_{32}^*B_{32} + 2B_{22}^*X_1^{-2}(I-D)B_{22}$ or

$$B_{32}^*B_{32} + B_{22}^*(I - 2X_1^{-2}(I-D))B_{22} = D - \frac{1}{2}X_1^{-2}. \tag{7}$$

Since $X_1 = g(D)$, a simple computation involving the expression of $g$ in (1) yields that $I - 2X_1^{-2}(I-D) \geqslant 0$. Hence (7) gives $D \geqslant X_1^{-2}/2 = g(D)^{-2}/2$. Again, from the expression of $g$ in (1), we derive that $D \geqslant I/2$ and thus

$$X_1 = g(D) = \frac{1}{\sqrt{2}}D^{-1/2}. \tag{8}$$

It follows from (7) that $B_{22}^*(I - 4D(I-D))B_{22} = 0$, which is the same as

$$0 = B_{22}^*(I - 4D + 4D^2)B_{22} = B_{22}^*(I-2D)^2B_{22}.$$

We thus obtain $(I-2D)B_{22} = 0$ or $DB_{22} = B_{22}/2$ as asserted.

To prove $DB_{23} = B_{23}/2$, we use (5) to derive that

$$\begin{aligned} 1 &= |B_{13}|^2 + 2B_{23}^*(I-D)^{1/2}X_1^{-2}(I-D)^{1/2}B_{23} \\ &= |B_{13}|^2 + 2B_{23}^*X_1^{-2}(I-D)B_{23}. \end{aligned}$$

Since $B$ is a contraction, we have $|B_{13}|^2 + B_{23}^*B_{23} \leqslant 1$. These two together yield $1 \leqslant 1 - B_{23}^*B_{23} + 2B_{23}^*X_1^{-2}(I-D)B_{23}$ or $B_{23}^*(I - 2X_1^{-2}(I-D))B_{23} \leqslant 0$. Since $I - 2X_1^{-2}(I-D) \geqslant 0$ as was noted before, we obtain $B_{23}^*(I - 2X_1^{-2}(I-D))B_{23} = 0$ and thus

$$0 = B_{23}^*(I - 4D(I-D))B_{23} = B_{23}^*(I-2D)^2B_{23}$$

by (8). Therefore, $(I-2D)B_{23} = 0$ or $DB_{23} = B_{23}/2$ as required.

From $DB_{22} = B_{22}/2$ and $DB_{23} = B_{23}/2$, we have $(I-D)B_{22} = B_{22}/2$ and $(I-D)B_{23} = B_{23}/2$ and thus $(I-D)^{1/2}B_{22} = B_{22}/\sqrt{2}$ and $(I-D)^{1/2}B_{23} = B_{23}/\sqrt{2}$. It follows from (3) that

$$A = \begin{bmatrix} 0 & 2B_{12} & 2B_{13} \\ 0 & \sqrt{2}B_{22} & \sqrt{2}B_{23} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{on } M = H_1 \oplus H_2 \oplus H_3. \tag{9}$$

On the other hand, since $M = \bigvee\{A^nx : n \geqslant 0\}$ and $H_2 = M \ominus (\bigvee\{x, y\})$, we have $H_2 = \bigvee\{P_2A^nx : n \geqslant 1\}$, where $P_2$ denotes the (orthogonal) projection from $M$ onto $H_2$. A simple computation with (9) shows that $P_2A^nx = (\sqrt{2}B_{22})^{n-1}(\sqrt{2}B_{23})x$ for all $n \geqslant 1$. Therefore,

$$\begin{aligned} D(P_2A^nx) &= D\left(\sqrt{2}B_{22}\right)^{n-1}\left(\sqrt{2}B_{23}\right)x \\ &= \frac{1}{2}\left(\sqrt{2}B_{22}\right)^{n-1}\left(\sqrt{2}B_{23}\right)x = \frac{1}{2}P_2A^nx \end{aligned}$$

if $n \geqslant 2$, and

$$D(P_2Ax) = D\left(\sqrt{2}B_{23}\right)x = \frac{1}{2}\left(\sqrt{2}B_{23}\right)x = \frac{1}{2}P_2Ax.$$

These show that $D = I/2$ and hence $X_1 = D^{-1/2}/\sqrt{2} = I$ by (8) or $X = \sqrt{2} \oplus I \oplus (1/\sqrt{2})$.

Finally, we prove that $M$ is a reducing subspace of $T$. Since $f$ is inner with $f(0) = 0$, we have $w(f(T)) \leqslant 1$ (cf. [4, Theorem 4]). This, together with $\|f(T)x\| = 2$, yields that the subspace $K \equiv H_1 \oplus H_3$ reduces

$f(T)$ and $f(T)|K$ has the matrix representation $\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ relative to the orthonormal basis $\{y, x\}$ of $K$ (cf. Corollary 4 or [18]). In particular, this gives $f(T)^*x = 0$ and $f(T)^*y = 2x$. Now we repeat these with $T$ and $f$ replaced by $T^*$ and $\widetilde{f}$, where $\widetilde{f}$ is the inner function $\widetilde{f}(z) = \overline{f(\overline{z})}$, $|z| < 1$. Since $\widetilde{f}(T^*) = f(T)^*$, we have $w(\widetilde{f}(T^*)) \leqslant 1$ and $\|\widetilde{f}(T^*)y\| = 2$. Letting $\widetilde{M} = \bigvee\{T^{*n}y : n \geqslant 0\}$, we infer from what were proved before for $T$ and $f$ that $\widetilde{A} \equiv T^*|\widetilde{M} = \widetilde{X}\widetilde{C}\widetilde{X}^{-1}$ for some operator

$$\widetilde{C} = \begin{bmatrix} 0 & \widetilde{C}_{12} & \widetilde{C}_{13} \\ 0 & \widetilde{C}_{22} & \widetilde{C}_{23} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{on } \widetilde{M} = \widetilde{H}_1 \oplus \widetilde{H}_2 \oplus \widetilde{H}_3$$

($\widetilde{H}_1 = \bigvee\{\widetilde{f}(\widetilde{C})y\}$ and $\widetilde{H}_3 = \bigvee\{y\}$) which is unitarily equivalent to $S(\widetilde{\phi})$ ($\widetilde{\phi}(z) = z\widetilde{f}(z)$ on $\mathbb{D}$), and $\widetilde{X} = \sqrt{2} \oplus I \oplus (1/\sqrt{2})$ on $\widetilde{M} = \widetilde{H}_1 \oplus \widetilde{H}_2 \oplus \widetilde{H}_3$. We check that $\widetilde{A}$ is unitarily equivalent to $A^*$. Indeed, since $C^*$ is unitarily equivalent to $S(\widetilde{\phi})$ and the latter is in turn unitarily equivalent to $\widetilde{C}$, there is a unitary operator $U$ mapping $M$ onto $\widetilde{M}$ such that $UC^* = \widetilde{C}U$. In particular, we have $U(\ker C^*) = \ker \widetilde{C}$ and $U(\ker C) = \ker \widetilde{C}^*$. Note that

$$\widetilde{f}(\widetilde{C})y = \frac{1}{2}\widetilde{f}(\widetilde{A})y = \frac{1}{2}\widetilde{f}(T^*)y = \frac{1}{2}f(T)^*y = x$$

by the analogue of (2). Hence $\ker C^* = \ker \widetilde{C} = \bigvee\{x\}$ and also $\ker C = \ker \widetilde{C}^* = \bigvee\{y\}$. Therefore, $Ux = \lambda_1 x$ and $Uy = \lambda_2 y$ for some scalars $\lambda_1$ and $\lambda_2$ of modulus one. Thus $U$ is of the form

$$U = \begin{bmatrix} & & \lambda_1 \\ & U_1 & \\ \lambda_2 & & \end{bmatrix}$$

from $M = H_1 \oplus H_2 \oplus H_3$ to $\widetilde{M} = \widetilde{H}_1 \oplus \widetilde{H}_2 \oplus \widetilde{H}_3$ and hence

$$\begin{aligned} U^*\widetilde{A}U &= U^*\widetilde{X}\widetilde{C}\widetilde{X}^{-1}U \\ &= \begin{bmatrix} & & \overline{\lambda_2} \\ & U_1^* & \\ \overline{\lambda_1} & & \end{bmatrix} \begin{bmatrix} \sqrt{2} & & \\ & I & \\ & & 1/\sqrt{2} \end{bmatrix} \widetilde{C} \begin{bmatrix} 1/\sqrt{2} & & \\ & I & \\ & & \sqrt{2} \end{bmatrix} \begin{bmatrix} & & \lambda_1 \\ & U_1 & \\ \lambda_2 & & \end{bmatrix} \\ &= \begin{bmatrix} 1/\sqrt{2} & & \\ & I & \\ & & \sqrt{2} \end{bmatrix} \begin{bmatrix} & & \overline{\lambda_2} \\ & U_1^* & \\ \overline{\lambda_1} & & \end{bmatrix} \widetilde{C} \begin{bmatrix} & & \lambda_1 \\ & U_1 & \\ \lambda_2 & & \end{bmatrix} \begin{bmatrix} \sqrt{2} & & \\ & I & \\ & & 1/\sqrt{2} \end{bmatrix} \\ &= X^{-1}U^*\widetilde{C}UX = X^{-1}C^*X = A^*. \end{aligned}$$

Finally, we check that $\widetilde{M}$ is contained in $M$. This is because, for any $n \geqslant 0$, the equalities

$$\begin{aligned} \|T^{*n}y\| &= \|\widetilde{A}^n y\| = \|UA^{*n}U^*y\| \\ &= \|UA^{*n}(\overline{\lambda_2}y)\| = \|A^{*n}y\| = \|(T|M)^{*n}y\| \end{aligned}$$

hold, which yields that $T^{*n}y$ belongs to $M$. Similarly, we can show that $M \subseteq \widetilde{M}$. Hence $M = \widetilde{M}$ and $T^*M = T^*\widetilde{M} \subseteq \widetilde{M} = M$. Thus $M$ reduces $T$. This completes the proof. □

## Acknowledgments

We thank the (anonymous) referee for his suggestions which lead to considerable improvements in the exposition.

## References

[1] T. Ando, Structure of operators with numerical radius one, Acta Sci. Math. (Szeged), 34 (1973) 11–15.
[2] H. Bercovici, Operator Theory and Arithmetic in $H^\infty$, Amer. Math. Soc., Providence, 1988.

[3] C.A. Berger, Normal Dilations, Ph.D. Dissertation, Cornell Univ., 1963.
[4] C.A. Berger, J.G. Stampfli, Mapping theorems for the numerical range, Amer. J. Math. 89 (1967) 1047–1055.
[5] M.J. Crabb, The powers of an operator of numerical radius one, Michigan Math. J. 18 (1971) 253–256.
[6] M. Crouzeix, Bounds for analytical functions of matrices, Integral Equations Operator Theory 48 (2004) 461–477.
[7] S.W. Drury, Symbolic calculus of operators with unit numerical radius, Linear Algebra Appl. 428 (2008) 2061–2069.
[8] E. Durszt, On the structure of unitary $\rho$-dilations, Acta Sci. Math. (Szeged) 28 (1967) 299–304.
[9] H.-L. Gau, P.Y. Wu, Lucas' theorem refined, Linear and Multilinear Algebra 45 (1999) 359–373.
[10] K. Gustafson, D.K.M. Rao, The Field of Values of Linear Operators and Matrices, Springer, New York, 1997.
[11] P.R. Halmos, A Hilbert Space Problem Book, second ed., Springer, New York, 1982.
[12] T. Nakazi, K. Takahashi, Two-dimensional representations of uniform algebras, Proc. Amer. Math. Soc. 123 (1995) 2777–2784.
[13] K. Okubo, T. Ando, Constants related to operators of class $C_\rho$, Manuscripta Math. 16 (1975) 385–394.
[14] A. Rácz, Sur les transformations de classe $C_\rho$ dans l'espace de Hilbert, Acta Sci. Math. (Szeged) 28 (1967) 305–309.
[15] D. Sarason, Generalized interpolation in $H^\infty$, Trans. Amer. Math. Soc. 127 (1967) 179–203.
[16] B. Sz.-Nagy, C. Foiaş, Similitude des opérateurs de classe $C_\rho$ à des contractions, C.R. Acad. Sci. Paris série A 264 (1967) 1063–1065.
[17] B. Sz.-Nagy, C. Foiaş, Harmonic Analysis of Operators on Hilbert Space, North Holland, Amsterdam, 1970.
[18] J.P. Williams, T. Crimmins, On the numerical radius of a linear operator, Amer. Math. Monthly 74 (1967) 832–833.